

ACUPUNCTURE FOR ELBOW PAIN

Bandolier is always attracted to systematic reviews claiming to find strong evidence about a treatment effect. Strong evidence is not easy to find, but one expects a rock on which good practice can be built. When strong evidence and acupuncture are used in the same sentence, though, *Bandolier* begins to smell a rat. Not because of any bias, but because good evidence indicates that acupuncture does not work.

It should be an educational experience, then, to see how a review of acupuncture for elbow pain [1] stacks up against the requirements of quality, validity, and size (QVS). What does this mean?

- ◆ Quality: trials that are randomised and double blind, to avoid selection and observer bias, and where we know what happened to most of the subjects in the trial.
- ◆ Validity: trials that mimic clinical practice, or could be used in clinical practice, and with outcomes that make sense. For instance, in chronic disorders we want long-term, not short-term trials. We are not interested in small changes of marginal statistical significance ($p < 0.05$, say, or a 1 in 20 chance of being wrong), but changes that are large, useful, and statistically very significant ($p < 0.01$, a 1 in 100 chance of being wrong).
- ◆ Size: trials (or collections of trials) with large numbers of patients avoid being wrong because of the random play of chance. For instance, to be sure that a number needed to treat (NNT) of 2.5 is really between 2 and 3, we need results from about 500 patients. If that NNT is above 5, we need data from thousands of patients.

These are the criteria on which we should judge evidence. For it to be strong evidence, it has to meet all three criteria.

The review

The review sought randomised studies of patients with pain resulting from tennis elbow, with other descriptions, but essentially with pain originating from the common origin of the extensor tendon, and with needle acupuncture as the primary intervention. Excluded were other elbow problems, and patients concurrently receiving other treatments.

Results

Of 53 articles screened, the authors chose to include six. Results from all these six trials were combined in a qualitative "best evidence synthesis".

In this issue

Acupuncture for elbow pain	p. 1
Stroke rates fall	p. 2
Alcohol dependence treatments	p. 3
Brief interventions for alcohol problems	p. 5
Campaigns about alcohol and driving	p. 6
RCT of BNP	p. 7

Of these six included trials, one was not properly randomised.

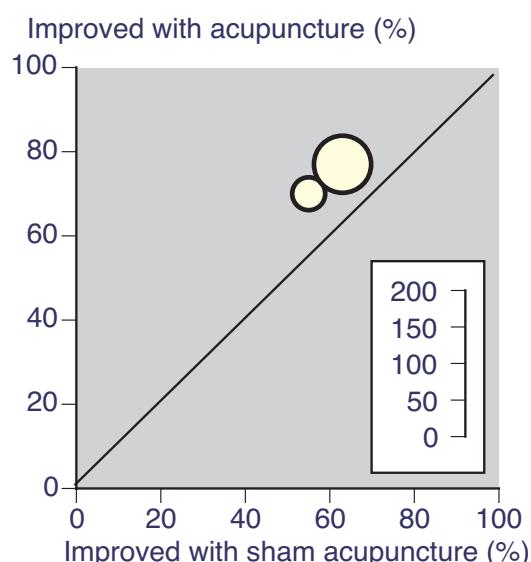
Of the remaining five trials, two were not double blind.

Of the remaining three, one had results only immediately after treatment (not much use in a chronic condition, and duration of the condition was about 10 months in some of the trials).

That left two randomised, double-blind trials, both of which reported results two or three months after treatment. Both compared real acupuncture with sham acupuncture (using different needle points, for instance). Both reported outcomes roughly equivalent to half pain relief, and the rates were similar (Figure 1).

There was no significant difference between real and sham acupuncture. The relative benefit was 1.2 (0.96 to 1.6).

Figure 1: Valid, randomised trials of acupuncture versus sham acupuncture for tennis elbow



Comment

This review failed QVS criteria. Six original trials soon became two small and useful trials, with valid outcomes, but no different from control. Even if there were an effect, it is so small that we would need large trials to be confident that it was there: our information is from only 123 patients.

And that is the best one might say. Reading the original trials is interesting. One included trial was published twice (without acknowledging the other), quoted odds ratios to six significant figures (52.2888 seemed slightly over-precise based on 48 patients), and used NNTs when results were

not statistically significant. That does not give much confidence in the quality of trial reporting.

So is there strong evidence for acupuncture in tennis elbow? No. Actually there is no evidence that it works, and what evidence we have suggests it does not work in any meaningful way. A pity, then, that this will be seized upon as a way of extracting money from wallets.

Reference:

- 1 KV Trinh et al. Acupuncture for the alleviation of lateral epicondyle pain: a review. *Rheumatology* 2004 (advance access 22 June, doi: 10.1093/rheumatology/keh247).

STROKE DOWN

Bandolier 125 examined the evidence that changes in both medical treatments and lifestyles had contributed to the reduction in coronary heart disease in industrialised countries over recent decades. A new study in Oxfordshire [1] paints a similar picture for stroke.

Study

What we have is two studies. In 1981-84 the Oxford Community Stroke Project examined stroke in people registered with family doctors in Oxfordshire. The Oxford Vascular Study repeated the exercise for 2003-4, using similar methods. Both studies went to enormous lengths to identify strokes through collaboration with family doctors, computerised codes in the hospital serving the population, visits to neurology wards, vascular imaging lists, eye hospital referrals, coroner's office registers, and others.

Patients were assessed as soon after the event as possible, with a standard clinical history and examination, with scanning in every case. Diagnosis and clinical subtyping in the more recent study were as close as possible to those in the earlier study.

The most recent measurement of blood pressure was recorded, and total cholesterol measured at assessment of event. Surviving patients were followed up for 12 months and a disability score calculated (*Bandolier* 124).

Results

Both studies involved populations of about 90,000 under the care of about 65 family doctors, and each study had several hundred strokes or transient ischaemic attacks. The population was stable over the period of the two studies, with little difference in racial mix. Organisation of health care was similar in practice, even if labels kept changing.

Over the 20 years between the studies there was a significant ageing of the population. The proportion aged less than 35 years fell by nearly 8%, with increases in all older age groups for 2004 over 1984 (Figure 1). The incidence of first stroke fell in almost all of these age groups (Figure 2), particularly in the age range of 55 to 85 years.

Significant reductions in stroke incidence for 2004 compared with 1981-4 occurred overall, in men and women, and in all patients aged less than 85 years. Significant reductions occurred in ischaemic and haemorrhagic stroke, but not subarachnoid haemorrhage. Strokes in 2004 were significantly less severe than those occurring in 1981-4.

Overall stroke incidence fell by about 30% between 1981-4 and 2004. On the basis of the demographic changes with an older population, extrapolating 1981-4 stroke incidence to 2004 would have predicted an increase in stroke incidence of about 30%.

Table 1: Percentage of pre-morbid treatments and risk factors in persons with incident stroke or TIA in 1984 or 2004

Treatment/risk factor	Incident stroke		Incident TIA	
	1981-84	2004	1981-84	2004
Anti-hypertensive	20	47	no data	46
Antiplatelet	3	34	5	38
Lipid lowering agent	0	11	0	22
Mean cholesterol over 6.0 mmol/L	58	30	67	36
Systolic BP over 150 mmHg	61	46	67	45
Diastolic BP over 85 mmHg	59	40	56	35
Current smoker	33	18	31	15

Figure 1: Changes in age-structure of population between 1984 and 2004

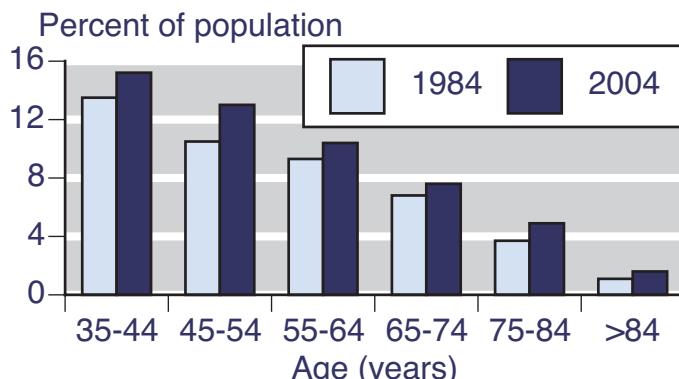
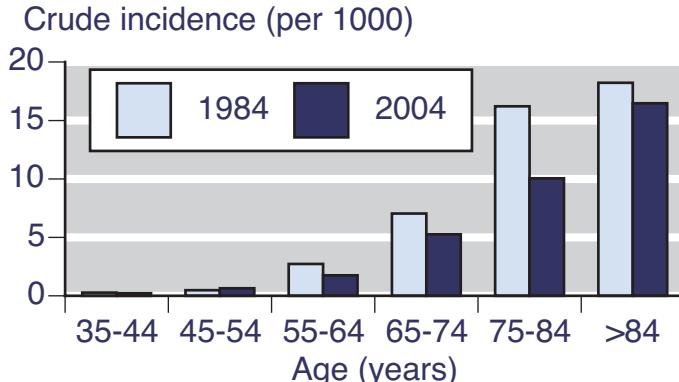


Figure 2: Changes in crude stroke incidence between 1984 and 2004



Treatments and risk factors

There were large changes in both pre-morbid treatments and risk factors for people with an incident stroke and those with incident transient ischaemic attack (Table 1). Compared with 1981-4, in 2004 many more patients were receiving anti-hypertensives, anti-platelets, and lipid lowering agents.

Risk factors were also lower. Fewer people had high cholesterol or blood pressure, or were current smokers. Compared with 1981-4, average cholesterol levels in those with a first event were lower by about 1 mmol/L, mean systolic blood pressure by about 10 mmHg, and mean diastolic blood pressure by about 6 mmHg.

Comment

This detailed study took an in-depth look at stroke incidence in Oxfordshire. Because of an ageing population, more strokes would have been expected. Fewer occurred. Both treatment of pre-morbid conditions, and improved risk factors contributed, but how much by each is not clear. Oxfordshire probably does as well as most places in implementation of preventative medicine, and many people have taken on board healthy living advice. Together these have resulted in reducing the number of strokes in the population of 2004 by about 55 per 100,000 people over what would have been expected without them. This is beginning to look like good value.

Reference:

- 1 PM Rothwell et al. Change in stroke incidence, case fatality, severity, and risk factors in Oxfordshire, UK from 1981 to 2004 (Oxford Vascular Study). *Lancet* 2004; 363: 1925-1933.

ALCOHOL DEPENDENCE TREATMENTS

Alcohol dependence is a bad thing. It has high impacts on individuals, their health, their family, and the economy. There are lots of suggestions for dealing with alcohol dependence, including drug treatments. A new systematic review [1] gives us a good idea of the efficacy of two of those treatments, acamprosate and naltrexone.

Systematic review

The review sought randomised studies of full, published papers comparing acamprosate or naltrexone with placebo or other control without medication. Studies had to be longer than two weeks. The period covered in searching four electronic databases was 1990-end 2002.

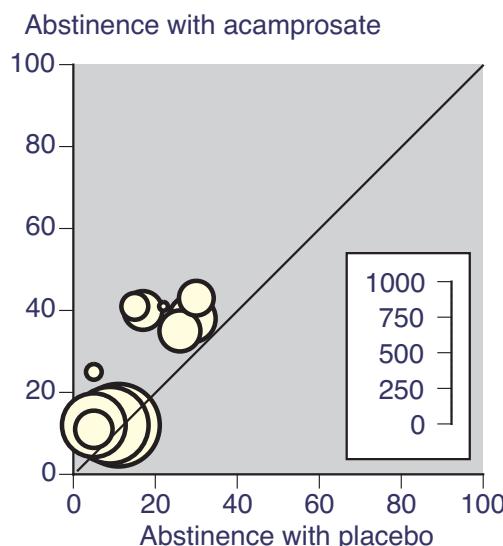
Results

Acamprosate

For acamprosate there were 13 placebo-controlled trials with about 4,000 subjects. Twelve were conducted in an ambulatory setting and were included in the review. All used standard definitions of alcohol dependence and all subjects had undergone a previous detoxification process. Most were double blind, and all had adequate quality scores. All but one reported an intention-to-treat result. Duration was three to 24 months, with most being six to 12 months. Most were definitely funded by the manufacturer of acamprosate; for four no information on funding source was available. The dose of acamprosate was most often set by body weight, being about 2,000 mg daily in those over 60 kg, and 1,300 mg a day in those below 60 kg.

Abstinence rates (percentage of patients completing the study without ingesting alcohol) varied (Figure 1), but with acamprosate were consistently higher than placebo. The overall abstinence rate was 23% with acamprosate and 15% with placebo, giving a number needed to treat to generate one more abstinent patient of 12 (95% CI 9 to 17; Table 1).

Figure 1: Abstinence rates with acamprosate and placebo



Acamprosate also had a small but significant increase in treatment compliance (Table 1). Gastrointestinal adverse events were higher with acamprosate (17% versus 11% with placebo, producing an NNH of 17; 12 to 27). Adverse event withdrawals were low, and not significantly higher with acamprosate than with placebo.

Naltrexone

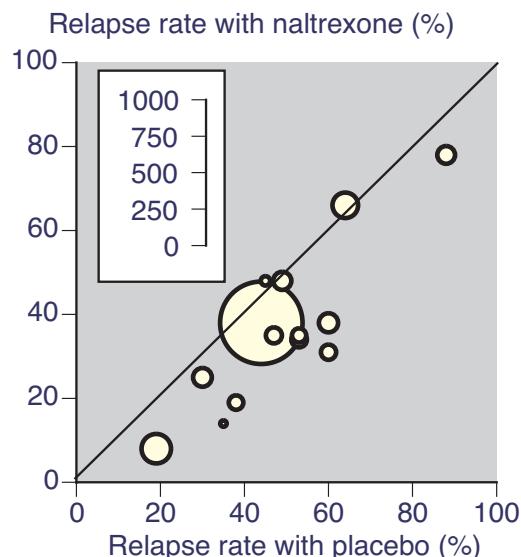
For naltrexone there were 19 trials, of which 12 were randomised and double blind, and from which information was taken; all had adequate quality scores. Information was available from over 2,000 subjects. One other (large) study was double-blind, but did not specify that it was randomised.

Studies used standard definitions of alcohol dependence, with subjects having undergone a previous detoxification. All were in an ambulatory setting. Duration was three to 18 months, with most studies of three or six months. Most were funded from charitable or government sources, with six funded in whole or part by a manufacturer. The dose of naltrexone was generally 50 mg daily.

There was no consistent reduction in abstinence rate with naltrexone. The abstinence rate with naltrexone of 35% was not significantly higher than the 30% with placebo (Table 1). There was a significant reduction in relapse rate to 37% with naltrexone compared with 48% with placebo (Figure 2), with an NNT of 10 (7 to 16).

Naltrexone produced a higher rate of gastrointestinal adverse events than placebo, with an NNH of 13 (9 to 21; Table 1). It also produced a higher rate of neuropsychiatric adverse events, with an NNH of 19 (11 to 73), but not serious neuropsychiatric adverse events. Adverse event discontinuation was significantly higher with naltrexone, with an NNH of 15 (11 to 22).

Figure 2: Relapse rate with naltrexone and placebo



Comment

Here we have two therapies, assessed in good trials which were of adequate quality, were valid, and with a reasonable number of patients in the combined trials. The message is reasonably clear, with acamprosate clearly better than placebo, with few adverse events, and naltrexone of questionable benefit and with more adverse events.

In randomised trials of alcohol treatment, some untreated participants (waiting list, no treatment, placebo) will give up or reduce their drinking by themselves. A review of untreated individuals [2] showed an average abstinence rate of 21% in 17 studies, with a mean level of consumption of 31 drinks a week, about six per week lower than the baseline of 37 drinks per week in 19 studies. This gives an indication of what can be achieved by doing nothing, and provides a benchmark for active treatments.

Table 1: Outcomes for benefit and adverse events with acamprosate and naltrexone compared with placebo in randomised trials

Outcome	Number of		Event rate (%) with		Relative benefit/risk (95% CI)	NNT/NNH (95% CI)
	Trials	Patients	Treatment	Placebo		
Acamprosate						
Abstinence	11	3322	23	15	1.6 (1.4 to 1.9)	12 (9 to 17)
Compliance	12	3959	53	47	1.1 (1.05 to 1.2)	18 (12 to 43)
Gastrointestinal adverse events	10	3425	17	11	1.5 (1.3 to 1.8)	17 (12 to 27)
Adverse event discontinuation	9	2697	2.8	2.2	1.3 (0.8 to 2.1)	
Naltrexone						
Abstinence	10	1077	35	30	1.2 (0.98 to 1.4)	
Relapse	14	2071	37	48	0.78 (0.71 to 0.86)	10 (7 to 16)
Gastrointestinal adverse events	17	2564	24	16	1.5 (1.3 to 1.8)	13 (9 to 21)
Neuropsychiatric	17	2564	43	37	1.14 (1.03 to 1.26)	19 (11 to 73)
Adverse event discontinuation	11	1892	10	3.6	2.9 (1.8 to 4.2)	15 (11 to 22)

NNH is **bold** and shaded

Which is the best choice? There may be differences in opinion about outcomes and their utility, but it would be difficult to argue that naltrexone was better than acamprosate.

In the past some would quibble. They might argue that these are indirect comparisons, and we could trust only a direct head-to-head comparison. There is one such. It had 157 patients, and concluded that naltrexone was better than acamprosate, though not by much. *Bandolier* 110 informed us that indirect comparison fulfilling quality, validity, and size criteria were at least as good, if not better, than direct comparisons lacking size.

The main question, though, may be what outcome you want. If it is abstinence, acamprosate clearly wins. If it is controlling consumption, rather than achieving abstinence, then naltrexone may have a role.

References:

- 1 B Carmen et al. Efficacy and safety of naltrexone and acamprosate in the treatment of alcohol dependence: a systematic review. *Addiction* 2004; 99: 811-828.
- 2 A Moyer, JW Finney. Outcomes for untreated individuals involved in randomized trials of alcohol treatment. *Journal of Substance Abuse Treatment* 2002; 23: 247-252.

BRIEF INTERVENTIONS FOR ALCOHOL PROBLEMS

While for some of us the problem with alcohol is not getting enough, for a distressing number of people the problem is drinking too much. There are, as always, gradations, but alcohol is a problem for a significant minority of people. Some recognise the problem and want to control their intake or stop. Others do not, or will not, recognise it as a problem.

One way of trying to deal with alcohol problems is that of a brief intervention, often by a primary care physician. Those brief interventions have all sorts of designations, and what is a brief intervention for one is an extended intervention for another. Then there is the difference between opportunistic interventions (often in primary care aimed at people with less severe problems), and those interventions that may be more targeted (often in hospital in people with more severe problems).

Two reviews [1, 2] give hope that brief interventions can be beneficial in some settings. The first [1] also has an excellent review of the issues of the interventions themselves and the populations in which they are used.

Alcohol problems

Three sources were used for studies in the first review [1], a specialised database, literature searches, and reviews and primary report citations. Brief intervention was defined as an intervention of no more than four sessions, though the amount of contact time is frequently difficult to estimate from reports. Studies were classified according to the type of comparison (brief intervention versus control, or a more

extended intervention), and the type of patient population (treatment-seeking versus non-treatment-seeking).

The outcomes examined were alcohol consumption, and a composite measure of multiple different types of outcome, 12 in all, that included alcohol consumption, abstinence, frequency of intoxication, global ratings, and problems in life resulting from drinking.

Results

Though 56 investigations were included, there was no definition of types of study architecture used, nor was there any indication of which (if any) were randomised trials.

Alcohol consumption

In 34 studies comparing brief interventions to control conditions in non-treatment-seeking populations, brief interventions had a significant though small effect on alcohol consumption in trials with a follow up from less than three months to up to 12 months.

In 20 studies comparing brief interventions to extended treatment conditions in treatment-seeking populations, there was no difference found.

Composite outcomes

In 34 studies comparing brief interventions to control conditions in non-treatment-seeking populations, brief interventions had a significant though small effect on composite outcomes in trials with a follow up from less than three months to up to 12 months.

In 20 studies comparing brief interventions to extended treatment conditions in treatment-seeking populations, there was no difference found.

Mortality

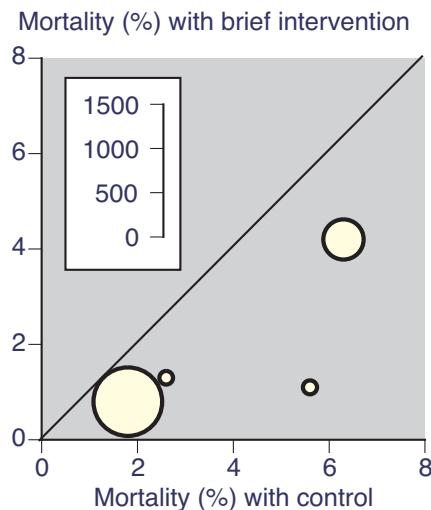
Whether brief interventions for problem drinking had any effect on mortality was examined in a second review [2]. Here studies included had to be randomised, compare a brief intervention to a no-intervention control in heavy drinkers, and have a long follow up with verified death rates. Studies were sought through consulting previous systematic reviews and meta-analyses, with additional searches of a number of electronic databases.

Results

Though there were 32 studies with 7,500 subjects meeting a broader set of inclusion criteria, only four with 1,540 subjects had verified death rates. Follow up in these four studies was between one and 10 years.

Death rates (Figure 1) were lower in the groups with a brief intervention compared with no intervention. Overall the death rate with brief intervention was 2.1%, compared with 3.3% with control. In terms of life years, there were 3 deaths per 1000 life years with brief intervention, compared with seven deaths per 1000 life years with control.

Figure 1: Brief interventions about alcohol, and mortality



The reduction was statistically significant with a relative risk of 0.5 (95% CI 0.3 to 0.96). For every 80-100 heavy drinkers given a brief intervention to stop or cut down their drinking, one fewer would die over the succeeding 1-10 years.

Comment

This is tricky territory for a number of reasons. First there is the intervention. What is it? A brief intervention defined for these reviews was up to four sessions of about 15 minutes or so, aimed at helping heavy drinkers stop or cut down. What is in a brief intervention? A variable feast, this, with a number of possible components. What is the outcome? Anything from reducing consumption to death. And whom are we targeting? There's the person who knows he or she has a problem, and wants help, and those not acknowledging the problem.

The problem we have with the first review is that we do not know the architecture or quality of the studies. Any result must therefore be taken with a pinch of salt, but, with that proviso, the results were fairly consistent. A brief intervention is better than nothing in those not seeking treatment. A brief intervention is probably as good as longer, more intensive, interventions in those who are seeking treatment.

What buttresses these results is the knowledge that in randomised trials with good ascertainment of outcomes, death rates were consistently lower with a brief intervention than with no intervention. Death is a hard outcome, and a reasonable proxy for other important outcomes.

While a brief intervention is not very effective, with an NNT of about 100, if an intervention can be constructed that is easy to do, it can make a difference, and may be well worth doing. Estimates of problem drinking suggest that significant portions of the population have alcohol problems.

References:

- 1 A Moyer et al. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. *Addiction* 2002 97: 279-292.
- 2 P Cuijpers et al. The effects on mortality of brief interventions for problem drinking: a meta-analysis. *Addiction* 2004 99: 839-845.

CAMPAIGNS ABOUT ALCOHOL AND DRIVING

Many years ago *Bandolier* saw a drunk driver take a corner at an insane speed and crash head-on into another car. The injuries were horrible, more to the crashed-into than the crasher. So campaigns in the media to tell people not to drink and drive have seemed like a good idea. Whether they work is another matter. A systematic review of studies [1] is both a compelling read, because it makes one think about what is involved in media campaigns, and provides compelling evidence that mass media campaigns work.

Background

This is not familiar territory for most of us, so a brief introduction to some aspects of media campaigns might help. It isn't as simple as one might think.

First is the content of the message. Is it going to be fear (pictures of bodies, drunk drivers in handcuffs and irons, courts, jail, whatever)? Is it going to be sweet reason (nominated drivers, responsible citizenship)? Whatever the choice, content is crucial, and may have to alternate between fear and reason to maintain awareness over the longer term.

Second is message delivery. The message has to be of good quality for a start. Think of the old-fashioned public service messages with cheap graphics and voice-overs compared with modern campaigns that are indiscernible from advertisements for cars, soft drinks, or beer.

Then there is the issue of testing the message and delivery before the campaign starts. You might know what it means, but the people on the receiving end might have a different perception. The paper gives the example of a campaign for responsible drinking where a third of viewers thought it was an advertisement for beer. The message and delivery are likely to change depending on the target audience, and with time.

Systematic review

For inclusion, studies had to be primary research published in English up to end-2001, related to mass-media interventions, have objective data on one or more outcomes related to alcohol-impaired driving, and meet minimum quality criteria for study design and execution. This last was a set of criteria laid out for evidence-based community preventive services.

Studies were only accepted if levels of other alcohol-impaired driving activities did not change substantially, or where statistical models accounted for any changes that did occur.

Studies were classified by whether they concentrated on legal consequences, or social and health consequences of drinking and driving. Outcomes used were fatalities or injuries related to alcohol-impaired driving, or people driving over the legal limit.

Results

There were six studies between 1975 and 1998, from Australia, New Zealand, and USA, with follow up from two to 37 months. A variety of interventions were used, and most were before-after studies, as well as time series with concurrent comparisons. All were evaluated in areas where actual enforcement levels during the campaign were at approximately the same level as before the campaign.

Two studies reported the proportion of drivers who were over the legal limit. Both had substantial reductions, of 30% and 37% in the number of drivers over the limit, and one had a very large sample base.

Four studies reported all crashes, or fatal, injury, or alcohol-related crashes where the driver had been drinking. All showed reductions, of between 6 and 18%.

Two studies also reported on cost-benefit analyses. One Australian study estimated costs of the campaign were (in US\$ adjusted to 1997) \$400,000 for advertising development, supporting media, media placement and concept research. Savings from medical costs were \$3.2 million per month, with total cost savings (medical, productivity, pain and suffering, property damage) of £8.3 million per month. A US study estimated total costs of a campaign to be \$800,000, with savings of \$7 million.

Comment

Preventative campaigns tend to be sporadic. For driving, in the UK we have a Christmas and New Year campaign, but little for the rest of the year. Perhaps we should do this more often, because the evidence is that it works. This may be a different sort of evidence from that we are used to, but one way of reducing medical costs is to advertise effectively against drinking and driving.

One additional thought is that this paper, while in an area unfamiliar to most of us, talks about evidence in ways we have come to expect. Quality, validity, and size are themes that run throughout. That is what makes the result believable.

A second additional thought is why limit this to drink driving? Why not ask, and try and answer, questions about getting over messages about healthy living, vegetables, fish, obesity, exercise, or sexually transmitted infections?

This paper opens a door on delivery of messages on health. It directly addresses some fundamental themes of how to make such media campaigns work effectively, and about how to assess their effects. One might even contemplate a future in which evidence-based advertising on good health vied with salt and fat rich snacks and prepared meals. Now there's a thought.

References:

- 1 RW Elder et al. Effectiveness of mass media campaigns for reducing drinking and driving and alcohol-involved crashes. A systematic review. American Journal of Preventive Medicine 2004 27: 57-65.

RCT OF BNP

Bandolier 121 examined the evidence for the use of B-type natriuretic peptide in the diagnosis of heart failure. On the evidence then available, BNP was probably going to be useful. The challenge was to work out how. The obvious thing to do was to perform a randomised trial of using the test in primary care, with the outcome of improved diagnosis. That trial has been done [1], and confirms the utility of the test.

Randomised trial

The trial, conducted in New Zealand, enrolled patients aged 40 years or more presenting to family doctors with symptoms of dyspnoea and/or oedema of recent origin. They could have comorbid conditions and be receiving any treatment. Those who needed urgent admission to hospital were excluded.

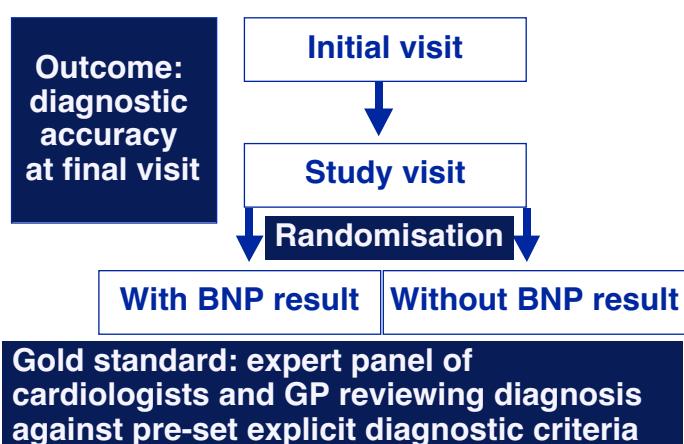
At the initial visit the family doctor recorded whether or not heart failure was suspected on history and examination. The patients were then booked to attend a study visit (Figure 1) in which they were assessed clinically by a cardiologist, with chest X-ray and ECG were conducted, and blood taken for BNP tests. A panel of three cardiologists and one family doctor used rigorous application of standard pre-defined criteria to make a gold-standard diagnosis of whether heart failure was present or not. This panel did not have access to BNP results, and was independent of study procedures.

Patients were then properly randomised to one of two groups (Figure 1). In one the family doctor had the results of the BNP faxed, with a standardised interpretive comment (Box). The other group had a laboratory sheet faxed informing the family doctor that the results of the BNP test were not available. The unit of randomisation was the patient.

Their family doctor, with or without the BNP result, then reviewed patients and made a final diagnosis. All family doctors had previously received a standardised 30-minute education session on interpretation of BNP results.

The outcome was the accuracy of the diagnosis made by the family doctor at the second visit.

Figure 1: Design of RCT to evaluate BNP test for diagnosis of heart failure in primary care



Standard interpretive advice with BNP results

- Normal range in healthy subjects BNP <50 pmol/L
- BNP >150 pmol/L strongly suggests heart failure in a newly symptomatic patient, but all clinical information should be taken into account.
- In between these levels heart failure is still possible, but all clinical information should be taken into account. BNP may be elevated by renal impairment, atrial fibrillation, LVH, COPD, after myocardial infarction, in the elderly and by treatment with beta-blockers or digoxin. BNP may be decreased by hypothyroidism, treatment with diuretics, vasodilators, and ACE inhibitors.

EDITORS

Andrew Moore
Pain Relief Unit
The Churchill, Oxford OX3 7LJ
Editorial office: 01865 226132
Editorial fax: 01865 226978
Email: bandolier@pru.ox.ac.uk
Internet: www.ebandolier.com
ISSN 1353-9906

Henry McQuay

the large differences at presentation between those patients with and without heart failure for parameters other than BNP might be used to create even more accurate diagnostic algorithms, particularly rule-out algorithms.

This is important because the majority of patients referred for heart failure don't have it. In this study an initial diagnosis of heart failure was made in 215 patients: only 77 were correct. Reducing unnecessary referrals should bring big benefits where hospital capacity is limited, or waiting times long.

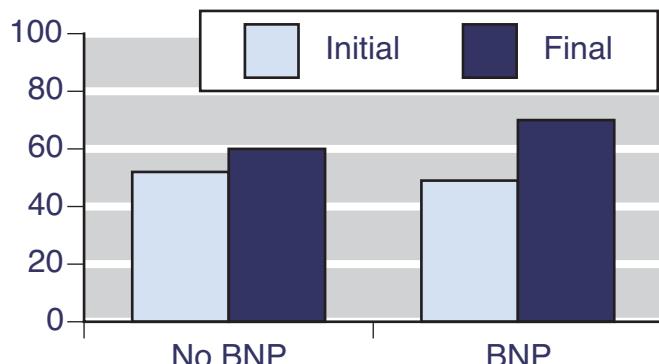
The study was supported by two New Zealand bodies, government and charity. The results will benefit patients, professionals and systems. Manufacturers of BNP testing kits will also benefit. Why is it, then, that diagnostic manufacturers do not undertake useful research to prove that their products really do make a difference?

Reference:

- 1 SP Wright et al. Plasma amino-terminal pro-brain natriuretic peptide and accuracy of heart-failure diagnosis in primary care. A randomized, controlled trial. *Journal of the American College of Cardiology* 2003 42: 1793-1800.

Figure 2: Overall percentage of correct diagnosis of heart failure by family doctor at initial and final visit, with and without BNP result

Percent correct diagnosis



Results

There were about 150 patients in each group, with an average age of about 70 years, and about two thirds were women. Comorbid conditions of hypertension, history of myocardial infarction, diabetes, asthma, and COPD were common. The two groups were well matched at baseline.

Panel diagnosis of heart failure was made in 77 patients, and 228 were judged not to have heart failure. Plasma BNP levels in those with heart failure averaged 290 pmol/L, compared with 61 pmol/L in those without heart failure.

At the initial visit the diagnostic accuracy of the family doctor compared with the panel judgement was about 50% in both groups (Figure 2). At the second visit, family doctors without BNP results improved their diagnostic accuracy by 8%, with an overall accuracy rate of 60%. With BNP results they improved their diagnostic accuracy by 21%, to 70%.

Almost all of the improvement in both groups came from correctly ruling out heart failure (Table 1). With or without BNP results, family doctors were good at diagnosing heart failure when it was actually present. They had more difficulty in ruling out heart failure when it was not present. The difference in diagnostic accuracy with BNP meant that about 11% more patients with symptoms of heart failure would be diagnosed correctly. The number needed to diagnose was about nine.

Comment

Here we have the missing link concerning the utility of BNP in primary care. On the basis of this well-conducted study, it looks worthwhile. One other interesting thought is that

Table 1: Diagnosis of heart failure or no heart failure by panel of cardiologists and family doctor versus family doctor, in patients with and without BNP result available

Diagnosis	With BNP			Without BNP		
	Panel	Family doctor	Percent correct	Panel	Family doctor	Percent correct
Heart failure	43	39	91	34	31	91
Not heart failure	109	67	61	119	60	50